DOCUMENT RESUME

ED 148 879 TM 006 769

AUTHOR	Rosenthal, Elsa J.; Beaton, Albert E.
TITLE	Annotated Bibliography.
INSTITUTION	College Entrance Examination Board, New York, N.Y.
SPONS AGENCY	Educational Testing Service, Princeton, N.J.
PUB DATE	May 76
NOTE	14p.; For related documents, see TM 006 750-776 and TM 006 805; For complete report, see TM 006 516
AVAILABLE FROM	College Board Publication Orders, Box 2815, Princeton, New Jersey 08541 (Item No. 2517248, $2.00)
EDRS PRICE	MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS	*Academic Achievement; Achievement Tests; *Annotated Bibliographies; College Bound Students; *College Entrance Examinations; Educational Problems; *Educational Trends; Elementary Secondary Education; Performance Factors; *Scores; Student Ability; Testing Problems
IDENTIFIERS	*Scholastic Aptitude Test; *Test Score Decline

ABSTRACT

Thirty-two reports and publications pertaining to national declines in abilities and in test scores are cited, and most are annotated. These publications were issued between 1961 and 1976 and provided evidence considered by the Advisory Panel on the Scholastic Aptitude Test Score Decline. (Author/EVH)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED148879

Annotated Bibliography

Elsa J. Rosenthal
Albert E. Beaton
Educational Testing Service

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY
College Entrance Examination Board
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TM006 769

Prepared for the Advisory Panel on
the Scholastic Aptitude Test Score Decline,
jointly sponsored by the
College Board and Educational Testing Service

May 1976

This appendix is one of 27 that have been published individually to supplement On Further Examination: Report of the Advisory Panel on the Scholastic Aptitude Test Score Decline. The report itself is available at $4 from College Board Publication Orders, Box 2815, Princeton, New Jersey 08540. Additional copies of this appendix, at $2 per copy, as well as a price list for all the appendixes, can be obtained from the same address. Payment must accompany any order not submitted on an institutional purchase order.

Copyright © 1977 by College Entrance Examination Board. All rights reserved.

Printed in the United States of America.

American Council on Education, "National Norms for Entering College Freshmen -- Fall 1966." ACE Research Reports, Vol. 2, No. 1, 1967.

American Council on Education, "The American Freshman: National Norms for Fall 1972." ACE Research Reports, Vol. 7, No. 5, 1972.

The ACE has developed norms for entering college freshmen annually for many years and for many variables, including racial composition. For the freshman year beginning in the fall of 1966, 10.8 percent in public four-year colleges were either black or American Indian (ACE 1967, p. 22). In 1972, this figure had increased to 15.9% (ACE 1972, p. 33). In 1972, 2.0 percent of entering college freshmen in public four-year colleges reported that they were either Mexican American or Puerto Rican (these categories had not been included in the 1966 questionnaire). That such increases in minority enrollment were not limited to public institutions is also indicated. In 1966, 2.4 percent of students in private universities were black or American Indian. In 1972, this figure was 5.4 percent.

Angoff, W., "Why the SAT Scores are Going Down." English Journal, March 1975, pp. 10-11.

Angoff reviews some hypotheses about factors that may have contributed to the decline of SAT scores. He also discusses others that appear not to have had any influence on the drop in scores. The hypotheses fall into four categories: (1) that there has been a change in the abilities of the high school population; (2) that there has been a change in the college-going population; (3) that there has been a change in the population that takes the SAT and similar college-entrance tests; (4) that the change is an artificial result of test properties and testing techniques.

Armbruster, Frank, The U.S. Primary and Secondary Education Process. With contributions by Paul J. Bracken and the assistance of Joyce Lind. Final Draft. Croton-on-Hudson, N.Y.: Hudson Institute, 1975.

This study was conducted as part of a project of United States social policy studies conducted under a grant from the U.S. Office of Economic Opportunity. Particularly relevant to the concerns of the Advisory Panel on Score Decline is an appendix citing test results from 22 annual state testing programs, from the District of Columbia and several large cities. The test scores show that the growth in academic achievement noted after World War II topped out during the 1960s. Many achievement indicators show this turning point occurring after the middle of the 1960s.

In a substudy by Paul J. Bracken, data are also cited from national renorming studies that point to significant declines in scores, especially in the upper grades. The Iowa Tests of Basic Skills were renormed in 1955, 1964, and 1970; the Comprehensive Tests of Basic Skills in 1973; and the Iowa Tests of Educational Development (ITED) in 1971. ITED reports also indicate that scores have continued to fall since 1971.

Armbruster rules out socioeconomic factors as major contributors to lower aptitude and achievement and concludes: "the problems of education are more likely to originate from forces within the schools than from outside factors. . . Perhaps a contributing factor was the same one which may have caused academic demands to be reduced: a loss of perspective by educators." (p.275) With the loss of perspective has come an erosion of traditional values and a decline in both social and academic discipline. Armbruster suggests that "our implicit trust in educators may not be well founded. . . Perhaps the most important first step is to

convince those with the final responsibility (policymakers) that a problem truly exists." (p. 278) "A glance at state and national data since the second half of the 1960s clearly shows that the entire primary and secondary school system was -- and is -- in jeopardy." (p. 280) He advocates that the federal government collect achievement data and use this information as a gauge in making financial grants to the schools.

Bullock, R., and Stern, J., "Estimated SAT Summary Statistics for High School Cohorts, 1967-1974." Memorandum, September 3, 1974. Princeton, N.J.: Educational Testing Service.

To answer the question of whether the number of junior scores of students who took the test again as seniors was a source of bias in the SAT "testing-year" means, special tables were developed in which only the senior scores of persons taking the SAT as juniors and seniors, i.e., "known" repeaters, were retained. The effect of removing these junior scores from the means for 1966-67 and 1973-74 was to reduce the mean decline in SAT-V from 27 points to 21 points, and the decline in SAT-M from 17 to 12 points.

Conference Board of the Mathematical Sciences, National Advisory Committee on Mathematical Education, Overview and Analysis of School Mathematics, Grades K-12. Prepared with the support of the National Science Foundation. Washington, D.C.: Conference Board of the Mathematical Sciences, 1975.

This study of the way mathematics is taught in the American public schools concludes that recent declines in mathematics test scores are not caused by faulty teaching but are a result of other factors that have caused similar nationwide declines in reading, social studies, and science test scores. The report reasons that "new math" courses are blameless because they are fundamentally sound and also because the spirit and substance of such courses have not been extensively implemented in the classrooms. Instead, because of declines in achievement levels in other school subjects, it is asserted: "Scholastic performance is sensitive to broad school and societal pressures that are little influenced by choice of content or teaching method in special subject areas."

Farr, R., Tuinman, J., and Rowls, M., Reading Achievement in the United States: Then and Now. Report prepared for the United States Office of Education by Educational Testing Service and Indiana University, 1975.

The investigators were concerned with whether the level of reading competency had declined over the years and sought to discover whether a data base was available to answer the question. They found that publishers kept renorming data confidential, they could find no data on readability levels of children's books, and they could not obtain records on reading performance of Army inductees. Census reports, too, were a blind alley, since the definitions of literacy were totally inadequate. Thus the data base for the study was the following: (1) extant research literature on studies of changes in reading performance; (2) a survey of 100 public school systems, to locate records of reading achievement change over a number of years; (3) a survey of the 50 states, to locate statewide reading achievement data. Despite the extremely inadequate extant data bases, the authors are willing to draw the following (tentative) conclusions: reading scores, as reported in the literature, appear to be higher for groups tested now rather than "then." School system achievement data show gradual improvement in reading competency for high school students, from

the beginning of the century up to 1965; there may, however, be a very slight decline in reading achievement after 1965.

Fasold, J., Hall, M., and Impara, J., Decline in Standardized Test Scores: A Widespread Phenomenon. Salem, Oregon: Oregon Department of Education, 1974.

A compilation of several articles on the theme of a general decline in the ability of high school students.

Flanagan, J. C., and Jung, S. M., Progress in Education: A Sample Survey (1960-1970). Palo Alto, California: American Institutes for Research, December 1971.

This study covered the reading comprehension of students in the Project TALENT sample. A 48 item reading comprehension test was administered in 1960 to more than 400,000 high school students representing a random sample of 1,353 different schools. Ten years later the identical test was administered to over 12,000 high school students in 134 schools, the sample being selected in the same way as the 1960 sample. Over this 10-year period, the mean score increased slightly from 30.81 to 31.25.

Foster, B. J., Statistics of Public Elementary and Secondary Day Schools, Fall 1972. Department of Health, Education, and Welfare Publication No. (OE) 73-11402. Washington, D.C.: United States Government Printing Office, 1973.

Harnischfeger, A., and Wiley, D., Achievement Test Score Decline: Do We Need to Worry? Chicago: CEMREL, 1975.

"Score, Test, and Pupil Evidence for Test Score Changes" (summary, pp. 69-72):

"For the past decade, nearly all reported test data show declines from grade five onward. The declines become more pronounced at higher grade levels. This pattern is obvious in all tested areas. ... A few test data first seemed to contradict these findings:

"The PSAT data of eleventh graders did not show a systematic decline. But we also found factors that explain the apparent nondecline. After correcting for scale drift, the PSAT data follow the general decline pattern from the mid-sixties to 1970 and from 1972 onward. The score increase in 1971 could be due to changes in the test-taking population, as in that year the PSAT was connected to the National Merit Scholarship Competition, resulting in a more selective test-taking group. On the whole, we conclude that the analysis of the apparently contradictory results shows that accounting for scale drift and change in tested student populations alters the PSAT score pattern to conform with the common mold.

"NAEP's finding of increasing Reading-Literacy performance for eleventh and twelfth graders (17-year-olds) contradicts all other test findings for this age group. Here we believe the key for explanation lies in the test contents and tasks. NAEP's items address basics far below the eleventh and twelfth grade curriculums, while the other tests are oriented toward more grade-level specific content. Typical grade-specific achievement tests and the NAEP assessment are not easy to relate. However, we should assert that the NAEP increase in Reading-Literacy performance is valid and important, even though it is not an achievement area for eleventh and twelfth graders and it may not ultimately contradict the findings of decline in more complex content.

"The changes in Science performance among secondary pupils are somewhat more puzzling... The college-bound group taking the ACT shows fluctuation in Science achievement over the period between 1965-1966 and 1973-74 of about 12 percent of a standard deviation even though there is no systematic trend. This fluctuation is smaller than the size of the comparable drops in English (27 percent) and Social Studies (34 percent), but is still sizable. The pattern for Science was a low period from 1966-67 to 1968-69 followed by a rise, a small drop, and a recent rise to previous highest levels. The same age group, in their performance on a very similar test, the ITED test of Natural Science, shows decline. Twelfth graders' performance, in the State of Iowa, showed a large decline between 1968 and 1970, with performance plateaus before and after. The decline, therefore, matches a period of rise in ACT performance. Twelfth graders on the NAEP also show large declines during an overlapping period (1969-73). The only known systematic discrepancy between these data sets is in the college-boundness of the ACT test takers. Perhaps science achievement has dropped precipitously among the noncollege bound, but not among those preparing for higher education.

"Another startling finding which we could not explain by means of characteristics of the tests or test takers is the large drop in verbal achievement scores for females.

"Although we have to conclude that test scores declined for all tested achievement areas for grades five through twelve over the past decade, we found no evidence for declines in lower grades (grades two, three, and four), but perhaps even slight increases in achievement scores. Our analyses of tests and test takers indicates that these findings are real, not artifacts....

"The School, a Search Arena for Causes of Achievement Declines" (summary, pp. 106-107)

"Features of children's school environments, analyzed as to their potential relevance as determiners of achievement test score declines, carry highly discrepant burdens of likely responsibility. Pupil mobility or desegregation seem not to have any explanatory power, as regional and state differences contradict the widespread character of the declines. Teaching staff characteristics, such as experience and education, are possible distal causes but existing data do not reveal any evidence that they play a relevant role in the current test score drops. Also, for some of the factors connected to school reorganization in the past decade, such as continuing high school consolidation and the crowded state of elementary schools following the postwar baby boom, we could not draw any straightforward conclusion of powerful determination of decreasing achievement test scores. This similarity arises from the simultaneity of their accompaniment by massive funding increases and the inconsistency of their likely impact with grade level differences in the decline. And we lack data showing changes in pupils' learning motivation.

"Evidence for other contextual factors seems to reveal potential contributory power in the explanation of test score decreases. One such factor is pupil absence rate, which has steadily increased over the past decade, resulting in smaller average amounts of schooling for pupils, but also burdening the teaching process considerably.

". . .But the strongest explanatory power seems to come from curricular changes. Our gross data indicate a considerable enrollment decline in

academic courses. Secondary pupils have been taking fewer courses in general English and mathematics. But also enrollment in typical college preparatory courses, such as algebra, first-year foreign languages, and physics, is decreasing...."

Hawes, G. R., "The Decline of the SATs." CHANGE, November 1972, pp. 17-19.
Average test score levels in college entrance tests appear to be declining, yet many teachers feel that students are not less able academically than their predecessors. In a technical report for the Carnegie Commission on Higher Education, Paul Taubman and Terence Wales found that the average mental ability of college students, as reflected in IQ test scores, was higher in the early 1960s than was that of college students in the 1920s and 1930s, in spite of greatly increased enrollments since World War II. Yet in recent years the decline of SAT scores has been noted, a striking case being that of Horace Greeley High School in the well-to-do suburb of Chappaqua, New York. The school has kept records of the quartile distributions of four-year cumulative course-grade averages of graduating classes for the last 23 years. The median in each of the last six years has varied only from 82 to 84, fluctuating between 80 and 83 in the previous 17 years. Yet SAT medians have steadily declined from 1962's 575 to 1972's 522.

Hilton, T. L., Rhett, H., Broudy, I. L., Bower, C., Carter, M. M., Creech, F. R., and Echternacht, G. J., The Base-Year Survey of the National Longitudinal Study of the High School Class of 1972. Final Report. Princeton, N.J.: Educational Testing Service, 1973.

Jencks, C., Inequality: A Reassessment of the Effect of Family and Schooling in America. New York: Basic Books, 1972.
Argues that school quality has little effect on achievement, drawing upon major surveys conducted since 1960: Census Bureau studies of social mobility and income distribution, Project TALENT's surveys of American high schools, the Equality of Educational Opportunity Survey, and many others.

Johnson, Charles E., Jr., "Changing Trends in College Enrollment." College and University, Fall 1973, pp. 35-41.
Documents the many changes in colleges in the recent past: in the college age population, enrollment rates, proportion attending two-year or four-year colleges, proportion attending public or private colleges, and college costs and financing.

Maeroff, G., "The Missing School Yardstick: Measuring Gains and Losses." New York Times, March 17, 1975.
A call for better systems of monitoring pupil progress. Also comments on the fact that science scores on the National Assessment of Educational Progress have "slipped" since the tests were first administered in 1969-70 (the readministration took place in 1972-73). The greatest decline was evidenced by 17 year old students.

Marco, G. L., and Stern, J., Investigation of the "Repeater" Hypothesis Concerning SAT Score Declines. Princeton, N.J.: Educational Testing Service, April 1974.
The decline in senior SAT-verbal means from 1967 to 1973 was studied after adjustments to senior means were made under three different assumptions

about the proportion of repeaters. The samples were three mutually exclusive groups: (1) April or May juniors who did not repeat the SAT in their senior year; (2) April or May juniors who repeated the SAT in November or December; (3) November and December seniors who did not take the SAT in April or May as juniors. The three assumptions were these: (1) all the nonrepeating juniors repeated the test in November or December of the senior year; (2) the percentage of April/May juniors who repeated the test in November and December was the same in 1967 -- i.e., 67.32 percent; (3) the percentage of seniors in November and December who were repeaters was the same as in 1968 -- i.e., 33.83 percent, the year in which the highest percentage of seniors were repeaters. Since not all the juniors would ever be expected to repeat the test in the senior year, the first assumption leads to an overestimate of the effect of repeaters, but the other two give more reasonable estimates. The results indicate the "repeater hypothesis" accounts for about three points of the 26-point SAT-verbal mean score decline of November and December seniors from 1967-73. The effect on the decline of all seniors or all juniors and seniors tested during an academic year would be even less.

McCandless, S. A., The SAT Score Decline and its Implications for College Admissions. Paper presented at the 1975 Western Regional Meeting of the College Entrance Examination Board, January 1975.

Several hypotheses that have been offered to explain the score decline are here considered in detail: (1) the possible invalidity of the test; (2) the decreased numbers of SAT "repeaters"; (3) the instability of the score scale; (4) changing test-taking patterns; (5) change in the college-going population (less academically able because of expansion of the candidate group); (6) regional change in SAT-taking and immediate college entry; and (7) progressive decline in developed reasoning ability among young Americans generally. None of the available data were found to be decisively supportive of any hypothesis. Concludes with a discussion of the implications for the college of planning to arrest or reverse a projected score decline for prospective applicants.

The article contains the following useful tables: (1) SAT Score Averages for 1972 and 1973 graduates; (2) SAT Score Averages, by Sex, for 1972, 1973, 1974 graduates; (3) SAT Score Averages by Grade Levels for the Testing Years 1965-66 through 1973-74; (4) SAT-Verbal Score Distributions, for Graduates of 1972 and 1973 Taking the Test; (5) SAT Score Averages for SAT-Taking Students from High School Graduating Classes, 1967-1974; (5b) for Males; (5c) for Females; (6) SAT Score Averages 1956-57 through 1973-74; (7a) Estimates of the Rate of High School Graduation, Immediate College Entry, and SAT-Taking among Both Sexes, 1959-74; (7b) for Males; (7c) for Females; (8) Estimated SAT Score Averages for All Immediate College Entrants, 1961 and 1968; (9a) Percentages of 12th-Grade Males Classified by SES, Ability, and 13th-Grade Enrollment 1960-61 and 1967-68; (9b) for Females; (10) Percentage Distributions of SAT-Takers by Region of Residence, 1959-68; (11) Percentage Distributions of Immediate College Entrants, by Region of Residence, 1963 and 1968; (12) Primary Families with Civilian Dependent Members, 18 to 24 Years Old and Percent Enrolled Full-time in College by Family Income: October 1967 to 1972; (13) Estimated SAT Score Averages for All Seniors, 1960-61 and 1967-68.

Mcdu, C. C., and Stern, J., The Stability of the SAT Score Scale. Research Bulletin RB-75-9. Princeton, N.J.: Educational Testing Service, March 1975.

The authors investigated the consistency of SAT scales from 1963-73. Two verbal and two mathematical sections that had been administered in the 1963 and 1966 administrations were readministered to a random sample of the 1973 SAT candidate group. The raw scores on 1963 and 1966 forms were placed on the reporting scale for a 1973 form. The raw-to-scaled score conversions derived in 1963 and 1966 for equivalent raw scores representing the same ability levels on these forms were compared with their corresponding scores on the 1973 reporting scale. The equating procedures used in the experimental equating were identical with those used in the operational equating of SAT scores during the period of the study. However, in the operational equating, several equatings intervened between the tests that were equated directly in this study. The results indicated that for both 1963 and 1966 equatings, the 1973 candidate group would have earned lower scores on both SAT-verbal and SAT-mathematical had the experimental equating results rather than the operational equating results been used in reporting the December 1973 scores. The average difference was 14 points for SAT-verbal and 17 points for SAT-mathematical scores. The study data do not provide an explanation for the discrepancy between the results of the experimental and operational equating over the period of the study.

Further research is suggested to determine the bias in different equating methods, the effect of various methods of reliability estimates on equating results, the effect of variations in the performance of succeeding generations of candidates on common equating sections, the effect of the choice of equating sections, and the effect of the choice of old forms to which each new form is equated.

Munday, Leo A., Declining Admissions Test Scores. Iowa City: The American College Testing Program, 1976.

Over the last 10 years the decline in test score averages has been experienced by both ACT and SAT, and the extent of the decline -- about 2 percent to 3 percent of a standard deviation per year -- has been similar for both programs. There is variation by subject matter field; the decline in ACT scores has been most marked in social studies, there has been no decline in natural science, and there has been some decline in English and mathematics. The decline ir test score averages has been accompanied by an increased variability in the academic achievement of students tested. Declining admissions test scores are a national phenomenon, occurring in all regions and in nearly all states. The percentage of ACT-tested high scoring students has remained about the same over the last five years, while the percentage of low scoring students has increased. There are marked sex differences. Overall, the scores of men have not declined appreciably over the last five years, but the scores of women have declined considerably. This trend has been accompanied by an increase in the percentage of women among ACT-tested students.

Two possible explanations for declining admissions test scores were examined. The first concerned who takes the test and the second, how well prepared they are. To deal with these questions, it was necessary to use data from state testing programs and the National Assessment since ACT data describe college-bound students and not high school students

generally. The results for two states showed that in recent years, different people have probably been taking the test, and that starting in the late 1960s and continuing to the present, high school students have likely been less well prepared than previously. In different years one factor likely has more influence than the other, and in other years they combine in their influence. If the limited data studied have applicability to the national scene, it is possible to conclude that in recent years the ACT score decline has been largely due to the changed pool of test-takers. For their part, colleges and universities would seem to have adapted to the declining admissions scores of their incoming students. (Author's summary)

"Newest National Norms Show -- Decline in Scores Continues." ACTivity, October 1975, 13 (4), 5.

The 1974-75 norms for ACT Assessment test scores show a continued decline, yet reported high school grades continue to rise. The 1974-75 mean composite score for the four tests that make up the test section of the ACT Assessment is 18.6 (based on a 1-36 scale), compared to 18.9 a year ago and 19.2 in 1972-73. The mean high school grade point average (based on a four-point scale for self-reported grades in subject areas covered in the ACT Assessment test section) is 2.91 in 1974-75, compared to 2.86 the previous year and 2.81 in 1972-73. Composite scores are higher for males (19.5 against 17.8 for females); females' grade point averages are higher than those of their male counterparts (3.00 for females, 2.81 for males). More junior are taking the tests. Of the 900,000 taking ACT Assessment during the past year, 23 percent were high school juniors. Last year they were 20 percent of the total and 18 percent in 1972-73.

Perry, D., and Swanson, E. O., "Decline in Minnesota College Aptitude Test Scores." In J. Fasold et al., eds., Decline in Standardized Test Scores: A Widespread Phenomenon. Salem, Oregon: Oregon Department of Education, 1974.

The Minnesota Statewide Testing Program annually offers a Scholastic Aptitude Test, measuring verbal comprehension and reasoning, to all Minnesota eleventh-grade students. In 1966-67, 98 percent of all juniors took the test, but in 1972-73, the figure dropped to 91 percent. Both the mean and median have been steadily declining since 1968-69. In that year the juniors had a mean of about 33.0; in 1972-73 the mean score was 31.0 on the same test form. The standard deviation was 13.3 in 1968-69 and 12.4 in 1972-73. (Minnesota mean scores for juniors tested in 1963-64 through 1965-66 and for juniors tested in 1966-67 through 1968-69 stayed essentially level.) The statistics are not sufficient to explain the decrease. Fewer students participate, but those who do not take the test are generally lower-ranking students whose participation would have further decreased scores. The decline is due to the decrease in frequency of high scores combined with increase of frequency of scores close to the average. Perhaps the most likely explanation of the decrease in average scores is that student attitudes and motivation have changed.

Peterson, I., "College Textbooks being Simplified to Meet Needs of the Poor Reader." New York Times, November 7, 1974.

Textbook publishers have recently lowered the reading levels of college textbooks. Most of the pressure for improved "readability" came from community colleges that have had hugely expanding populations. There

was, however, evidence of sufficient trouble with reading comprehension at four-year colleges to persuade the publishers to produce new readability formulas.

Schrader, W. B., Test Data as Social Indicators. Statistical Report SR-68-77. Princeton, N.J.: Educational Testing Service, September 1968.

This survey was undertaken to find out whether useful year-to-year comparisons of educational achievement in the U.S. could be made using existing test data. Tests were located that provided the necessary data for making the comparisons, and all the 186 specific comparisons of earlier with later performance were expressed in a common unit -- the percentile rank on the new norms of a student who scored at the 50th percentile rank on the old norms. Using the norms for the Iowa Tests of Basic Skills (grades 3-8), the Metropolitan Achievement Tests (grades 5-8), and the Stanford Achievement Test (grades 5-8), Schrader found slight gains for students in the lower grades, who would have been high school seniors in 1968, over same-aged groups, who would have been seniors in 1957, 1960, and 1963.

Schrader, W. B., and Hilton, T., Educational Attainment of American High School Seniors in 1960, 1965, and 1972: Feasibility study. Final Report. Project No. NIE-G-74-0050, 1975. Princeton, N.J.: Educational Testing Service, 1975.

Schrader and Hilton, after a study of the feasibility of equating the 1960 Project TALENT, the 1965 Educational Opportunities Survey (Coleman 1966) and the 1972 National Longitudinal Study of the Class of 1972 (Hilton et al. 1973), propose an analysis of mean change of these national samples over the period of the score decline on the SAT.

Schrader, W. B., and Jackson, R., Preliminary Report on the 1974 PSAT/NMSQT Norms Study. Memorandum dated March 24, 1975. Princeton, N.J.: Educational Testing Service.

Using data from PSAT/NMSQT administrations in 1960, 1966, and 1970, the authors found that PSAT mean verbal and mathematical scores had remained virtually constant over a 14-year period, even though there was a slight increase in the 1966 means. In 1960, the mean verbal score was 34.2, and in 1974, it was 34.3. In 1960, the mean mathematical score was 38.7, and in 1974 it was 38.6. The schools were randomly selected, but from schools who voluntarily cooperated. The sample was not completely random because all schools selected did not participate in the study. In 1960, only seven-tenths of the schools participated, and in 1974 only three-fifths of the selected schools participated. Accordingly, the findings may be questioned.

Start, K. B., "Thirty Years of Reading Standards in England." Educational Researcher, 1972, 1 (11), 8-9.

A 1970-71 national survey of England's 11- and 15-year-old children was undertaken to establish current reading standards and to relate them to similar surveys that had been conducted in 1938, 1948, 1952, 1956, 1961, and 1964. Tests were administered to a stratified random sample, with school type and size providing the strata and the within-cell sample fractions drawn by birthdays. Scores in 1948 were below the 1938 estimates but had risen steadily since. When the 1970-71 figures were added to those of the previous twenty years, the data indicated that reading

standards of the children leaving English primary schools "were no better now than they were in 1960 and could well be less than the standards in 1964." An official committee of inquiry has been established to investigate how the teaching of English, training of teachers, and monitoring of educational attainment may be improved.

Stewart, E. E., The Stability of the SAT-Verbal Score Scale. Research Bulletin RB-66-37, CEEB RDR-66-7. Princeton, N.J.: Educational Testing Service, July 1966.

Raw scores on each of four old forms of the SAT-Verbal test were placed on the reporting scale for a newer form. The new data were collected as part of the December 1963 administration; data on the old forms were retrieved from files for the April 1944, January 1948, March 1953, and February 1957 SAT administrations. Results: (1) For both the 1953 and 1957 forms, the experimentally derived scales and the reporting scales differed by less than five points in the vicinity of the mean, thus good enough to conclude that the SAT-Verbal scale remained highly stable during 1953-63. (2) Although the standard errors of both scales are relatively large, it may be concluded that the score of a candidate tested in 1963 and scoring in the vicinity of 500 might be 20-35 points higher than the score of a candidate of equal ability tested in 1948. Since the scale as it existed in 1953 was kept through 1963, the shift in the meaning of the scale must have occurred between 1948 and 1953. Scores on SAT-Verbal reported in April 1944 may not have had the same meaning as those of December 1963, but the methodological limitations inherent in the experimental equating of those two forms preclude any reliable quantitative assessment of whatever differences may exist between the scale of the early 1940s and that of the early 1960s.

Tavris, Carol. "The End of the IQ Slump." Psychology Today, April 1976, pp. 69-74.

Tavris summarizes Zajonc's arguments (see below): Today's low test scores are owing to (preceded by) children born close together in large families. The "IQ Slump" will be over around 1980 because the score declines directly reflect the increased family size of the postwar years and the students' later birth orders. We are already beginning to see an upswing in achievement in lower grades.

Thorndike, R. L., "Mr. Binet's Test 70 Years Later." Educational Researcher, May 1975, 4 (5), pp. 3-7.

Traces the history of the test since its development, about 1905, in France, through its American revisions in 1916 and 1960, and renormings in 1937 and 1972, giving a qualified defense of the test's unitary, overall appraisal of level of mental functioning. In the 1972 renorming, the shifts in distribution of IQs were "dramatic," especially the increase of scores for preschoolers; moderate gains were also noted for adolescents. The high average values were impressive because the sample in 1972 included, for the first time, ethnic minorities. This rise is interpreted as evidence of the increase in verbal and visual stimulation of young persons, especially those of preschool age. It is noted that the largest improvement has taken place in preschoolers' performance on nonverbal, pictorial, perceptual, and memory items, rather than on those that involve a good deal of semantic content. Thorndike suggests that our conception and use of Binet's test be modified so as to accord with current conditions in the modern egalitarian, pluralistic world.

Wilks, S. S., Scaling and Equating College Board Tests. Princeton, N.J.: Educational Testing Service, 1961.

Zajonc, R. B., "Family Configuration and Intelligence." Science, 16 April 1976, pp. 227-236.
Variations in aggregate intelligence scores are closely associated with variations in patterns of family configuration, and these aggregate family factors are deeply implicated in the declining SAT scores as a special case of a general phenomenon that manifests itself also in a variety of national, ethnic, regional, racial, and sex differences in intellectual test performance.

Several findings are pertinent: (1) Intellectual performance increases with decreasing family size. (2) Children born early in the sibship perform better on intelligence tests than later children when intervals between successive births are relatively short. (3) Long intersibling spacing appears to cancel the negative effects of birth order and in extreme cases to reverse them. (4) In general, long intervals enhance intellectual growth. (5) The adverse effects of short intervals are reflected in the typically low IQs of children of multiple births. (6) In the special case of only children, the benefits of a small family are apparently counteracted by the lack of opportunities to serve as teachers to younger children. (7) Last children suffer that handicap too. (8) Absence of a parent is associated with lower intellectual performance by the children. (9) Temporal changes in family patterns such as birthrates, average orders of births, intervals between children, and family size are reflected in temporal changes in aggregate measures of intellectual performance. (10) Differences in family patterns between different countries, between different regions of the same country, and between ethnic or racial groups are also associated with differences in aggregate intellectual performance. (11) Males and females differ in average birth order, and this difference is reflected in aggregate intellectual performance scores.

The large decline in SAT scores is surely influenced by the influences listed above. The decline in SAT scores (over 1/3 S. D. in 12 years) cannot be a function of changes in family configuration alone because it is considerably larger than would be expected on the basis of a simple extrapolation from four national samples (The Netherlands, 1965; United States 1965 NMSQT scores; France, 1973 scores; Scotland, 1947 Stanford-Binet scores). It is not claimed that the confluence model generates a unique interpretation of all these facts. For each of them one could probably supply another reasonable explanation. The intellectual deficit of twins could have a biological basis, for example, and the higher intelligence of twins who lost their co-twins may involve unknown genetic factors. The drop in SAT scores may be due to a general decline in intellectual interests... . Future research will shed light on these questions. At the moment, however, the confluence model has the advantage of parsimony. And because it makes rather specific predictions, it can be readily verified... . (from the author's summary)